BMJ Open

Can risk be predicted? An Umbrella Systematic Review of current Risk Prediction Models for Cardiovascular Diseases, Diabetes and Hypertension.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-030234
Article Type:	Research
Date Submitted by the Author:	05-Mar-2019
Complete List of Authors:	Lucaroni, Francesca; University of Rome Tor Vergata, Department of Biomedicine and Prevention Cicciarella Modica, Domenico; University of Rome Tor Vergata, Biomedicine and Prevention Macino, Mattia; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention Palombi, Leonardo; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, biomedicina e prevenzione Abbondanzieri, Alessio; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention Agosti, Giulia; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention Biondi, Giorgia; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention Morciano, Laura; University of Rome Tor Vergata, Department of Biomedicine and Prevention Vinci, Antonio; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention
Keywords:	Cardiovascular diseases, Hypertension < CARDIOLOGY, Diabetes, Risk Prediction Models

Can risk be predicted? An Umbrella Systematic Review of current Risk Prediction Models for Cardiovascular Diseases, Diabetes and Hypertension.

Prof. Leonardo Palombi, Department of Biomedicine and Prevention, University of Rome Tor Vergata
Via Montpellier, 1 – 00133 – Rome, Italy
e-mail: leonardo.palombi@gmail.com
Tel +39 3358436196

Domenico Cicciarella Modica, Mattia Macino, Antonio Vinci, Alessio Abbondanzieri, Giulia Agosti, Giorgia Biondi:
School of Hygiene and Preventive Medicine, University of Rome Tor Vergata – Rome, Italy

Francesca Lucaroni, Laura Morciano:
Department of Biomedicine and Prevention, University of Rome Tor Vergata – Rome, Italy

Word count: 5081

Keywords: *Cardiovascular diseases*; *Hypertension*; *Diabetes*; *Risk Prediction Models*;

Can risk be predicted? An Umbrella Systematic Review of current Risk Prediction Models for Cardiovascular Diseases, Diabetes and Hypertension.

Lucaroni F[2] & Cicciarella Modica D[1], Macino M[1], Palombi L[2*], Abbondanzieri A[1], Agosti G[1], Biondi G[1], Morciano L[2], Vinci A[1].

[1]School of Hygiene and Preventive Medicine, University of Rome Tor Vergata – Rome, Italy
[2] Department of Biomedicine and Prevention, University of Rome Tor Vergata – Rome, Italy

*: Corresponding author

Background: Cardiovascular diseases (CVDs), hypertension and diabetes represent one of the major causes of morbidity and mortality worldwide, and area leading cause of disability. Being preventable conditions, early identification is a public health crucial issue.
Several Risk Prediction Models (RPMs) have been investigated in recent decades, but no consensus on gold standard tools has been reached. Aim of this study is therefore to compare the effectiveness of available RPMs in terms of early recognition of CVDs, diabetes and hypertension, in order to assess which are the most effective tools in order to program preventive interventions.

Methods: An umbrella systematic review was performed, according to PRISMA guidelines, on available studies as at February 2018, searching three engines (MEDLINE/PubMed, Scopus and Cochrane Library). The studies included were systematic reviews or meta-analyses on healthy adults (18-65yo).No restrictions were applied on publication date, whereas inclusion was limited to articles published in English. AMSTAR tool was used for quality assessment; risk-of-bias evaluation was performed using the ROBIS tool.

Results: Of 3,206 studies evaluated, 15 met the inclusion criteria: six on diabetes, two on hypertension, seven on CVDs. Results were extremely variable among studies and no unequivocal comparison tool was applied to test RPMs effectiveness on diabetes and hypertension.

Discussion and conclusions: Although a large number of the studies focused on effectiveness of risk prediction models for CVDs, hypertension and diabetes, no unequivocal results were available. Certain evidence of clear effectiveness was available only for cardiovascular disease prediction: Framingham score, alone or in combination with ankle-brachial index, and QRISK score could be reasonably considered the gold standard for early identification of CVDs. Further efforts should not be concentrated on creating new scores, but rather on performing external validation of the available ones, possibly on target populations, with no available evidence for specific exposure risks.

PROSPERO SRRN: CRD42018088012

ARTICLE SUMMARY:

Strengths and limitations

- This is the most comprehensive umbrella systematic review on RPMs for CVDs, Hypertension and Diabetes up to date;
- Available studies, although apparently of medium-high quality, were based on primary studies of debatable quality, several of which lacking discrimination and calibration assessments;
- Publication bias and reporting bias could not be assessed.
- Heterogeneity too high for meta-analysis; results were therefore reported narratively.

INTRODUCTION

Cardiovascular diseases (CVDs), hypertension and diabetes represent a major health concern, throughout the world, from high- to low-income countries, as a silent epidemic responsible for millions of deaths every year.

CVDs, excluding hypertension, have a global prevalence rate of 6.6% and account for 17.6 million deaths a year, becoming the leading cause of death worldwide (1).

Hypertension alone is the leading preventable cause of premature death worldwide(2), being responsible for 7.5 million deaths globally.
In 2010, almost one adult in three (31.1%) had hypertension, although there was a significant gap in prevalence rates between high- and low-income countries: 28.5% (27.3%-29.7%) and 31.5% (30.2-32.9%) respectively worldwide.

Diabetes (both types 1 and 2) has a prevalence rate of 5.4% globally and it is responsible for 1.4 million deaths every year(3).
Besides being a significant cause of mortality worldwide, CVDs, hypertension and diabetes are also a leading cause of disability. They account together for almost 40% of disability-adjusted life years (DALYs): cardiovascular diseases alone are responsible for 32.3% of total DALYs, hypertension for 3.7% and diabetes for 2.4% (4).

The epidemiology of these three diseases contributes to explaining the substantial economic impact they have on national health services globally: in the United States, cardiovascular disease-related direct costs amount to approximately $444 billion (5), whereas the costs of diabetes are estimated as$327 billion and hypertension is associated within annual estimated cost of $51 billion (6), most of which (nearly $48 billion) represents direct medical expenses (7).
In recent decades, hypertension and diabetes have shown an increasing trend in terms of prevalence and mortality rates.It has been estimated that the prevalence of diabetes will continue to grow: 1 in 5 to 1 in 3 adults will be affected by 2050. The same is true for mortality rate trends: in the last fifteen years, diabetes-related deaths increased by 1%, and these data are expected to climb dramatically over the coming decades (8).
The prevalence of hypertension and the associated mortality rates have increased significantly in both men and women(9).By 2030, projections estimate an increase in prevalence to 44%(10).
Conversely,CVDs have significantlydecreased in terms of prevalence and mortality rate trends (11), including in those countries that had seen considerable increases until the beginning of the 21st century(12).

Cardiovascular diseases, hypertension and diabetes are strongly related to each other: diabetes is associated with an increased risk of CVDs, which is exaggerated by concomitant hypertension. These conditions also share the same pathogenic pathways, at both macroscopic and molecular level: oxidative stress, inflammation and fibrosis, causing microvascular and macrovascular complications of diabetes, which also lead to vascular remodelling and dysfunctions in hypertension (13).
Because of the high prevalence and mortality rates associated with CVDs, hypertension and diabetes, and their related direct and indirect costs, early identification of these three diseases is a crucial issue, in terms of obtaining significant savings in global health outcomes and economic expenditures.
A number of prediction models focused on these big three non-communicable diseases (NCDs) are currently available, but there is no consensus as to the gold standard tools to be utilised to enable early identification of people at risk.
The aim of this study is to evaluate risk prediction models (RPMs) for CVDs, diabetes and hypertension, in order to assess which are the most effective tools in recognising vulnerable people at an early stage, in particular among high-risk groups (e.g. workers, migrants) and, consequently, to achieve primary and secondary targeted preventive interventions.

METHODS

This umbrella systematic review was performed according to a protocol designed *a priori*, following the PRISMA guidelines (14).
The PubMed, Scopus and Cochrane databases were searched electronically on 19 February 2018, using combinations of the relevant medical subject heading (MeSH) terms, key words, and word variants for "risk prediction scores" and "cardiovascular disease", "diabetes" and "hypertension", as shown in Table 1. The search and selection criteria were restricted to the following: systematic reviews, with or without meta-analysis, as type of study; general population aged 18-65 with no major illness; the comparison of at least two risk prediction models; English language. For the PubMed database only, further specific filters were added: only articles on humans and with abstracts available. No restrictions were applied in terms of publication date for all three databases.
Reference lists of relevant articles and reviews were hand-searched for additional reports.
The study protocol was registered on the PROSPERO (*International prospective register of systematic reviews*) database (Registration number: CRD42018088012).

Table1 – Search strategy

Database	Search string(s)	Filters
Cochrane Library	risk (prediction OR assessment) (model OR score)	/
	diabetes risk prediction	
	cardiovascular risk prediction	
	hypertension risk prediction	
PubMed	(risk[Title/abstract] OR "risk assessment"[MeSH Terms]) AND predict*[Title/abstract] AND (model[Title/abstract] OR score[Title/abstract])	Systematic reviews
		Meta-analysis
		Abstract
		Humans
		English
Scopus	(risk AND prediction AND model AND (systematic AND review OR meta-analysis))	Review
		English
		Medic

For each database, two different authors independently screened article titles and abstracts: MM and AA for Scopus, AV and DCM for PubMed/Medline, AV and AA for Cochrane Library.
Disagreements were discussed by the authors and resolved by consensus or by recourse to a third author (FL).
Every article meeting eligibility criteria -Systematic Reviews/ Meta-Analyses on risk prediction models on CVDs, Diabetes and Hypertension evaluated, by comparison with other RPMs, in adults with no relevant illness- was considered for subsequent qualitative synthesis.
Duplicate records were identified and removed, as well as articles including exclusion criteria: any study carried out with the only purpose to develop a new RPM, to validate one, or to propose a diagnostic/prognostic tool, without any comparisons with other prediction models.
Studies were henceforth labelled for inclusion or exclusion.
The selection process described above is summarised by the PRISMA flow diagram shown in Figure 1.

Assessment of study quality

Quality assessment of individual studies was performed by applying the AMSTAR tool (15).According to their score, articles were classified into three groups: low (AMSTAR score <4), medium (AMSTAR score ≥ 4 and ≤ 7), and high quality (AMSTAR score ≥ 8).

Four authors (AA, DCM, MM and AV) independently assigned score. Disagreements were resolved by consensus or by discussion with a fifth author (FL). No reviews were excluded *ex-post* for quality reasons.
Synthesis of the selected reviews was performed according to the guidelines for umbrella review from the Joanna Briggs Institute (16).
A qualitative (narrative) synthesis was performed, comprising the following variables: number of studies included, intervention type, comparison, performance indexes (AUC, C STAT., D-STAT), validation, outcomes (i.e. incidence, prevalence, mortality), quality assessment score.

Risk of bias in individual studies

Risk of bias in individual studies was independently evaluated, using the ROBIS tool (17), by four authors (AA, DCM, MM and AV). Any disagreements were resolved by consensus or by discussion with a fifth author (FL).

Patient and Public Involvement
As the study design was a systematic review, Patients and Public were not involved.

RESULTS

As shown in Figure 1, 3,206 studies were identified through a search of the electronic databases.
After title screening, 3,058 studies were excluded because they did not meet the eligibility criteria and nine duplicate articles were removed. Of the 139 studies passing the first evaluation stage, 103 studies were excluded because the topic was not pertinent, sample characteristics were inadequate (i.e. some articles included non-healthy populations), the article type did not meettheinclusion criteria (i.e. some studies were not systematic reviews), or there was no comparison between models.
Following reading of the full text of the 36remaining studies, only 12 were deemed relevant for inclusion. 24 studies were excluded because of a lack of comparison model, related but non-pertinent topic or wrong article type.
Three additional studies were included after consulting reference lists of relevant articles and reviews293225.
An overall number of 15 studies met the eligibility criteria and were included in qualitative synthesis.

Quality assessment of the studies

All of the studies were of high-moderate quality, according to the AMSTAR tool, with a mean score of 7.14, (range 5-11): 8 out of 15 (53.4%) were medium quality (AMSTAR ≥ 4 and ≤ 7), and 7out of 15 (46.6%) obtained a score over 8 and were categorised as high quality.
The results of quality assessment have been summarised in graph form, as shown in Figure 2.

Risk of bias within studies

The results of risk-of-bias evaluation are shown in Table 2 and in Figure 3.
According to the ROBIS tool, only eight articles out of fifteen (53.4%) had a low risk of bias.
Two articles had a high risk of bias due to the eligibility criteria: search limitation on English language. Eight articles had a high risk of bias due to the identification and selection process: the most common source of bias was the search limitation to a single database. Nine articles had a high risk of bias due to data collection and study appraisal, in particular because of the lack of formal appraisal tools. Finally, one article had a high risk of bias due to the synthesis and identification process, mainly due to significant heterogeneity of primary sources, and it was subsequently impossible to conduct a meta-analysis.

Table2 - Risk of bias

Review	Phase 2				Phase 3
	1. Study eligibility criteria	2. Identification and selection of the studies	3. Data collection and study appraisal	4. Synthesis and findings	Risk of bias in the systematic reviews included
Abbasi	☺	☹	☹	☺	☹
Barber	☺	☹	☹	☺	☹
Beswick	☺	☺	☺	☺	☺
Collins	☺	☹	☹	☺	☺
Cortes-Bergoderi	☺	☹	☹	☺	☹
Damen	☹	☹	☺	☺	☺
Echouffo-Tcheugui 2013	☺	☺	☹	☺	☺
Echouffo-Tcheugui 2015	☺	?	☺	☺	?
Fowkes	☺	☺	☹	?	?
Hu	?	?	?	☹	?
Noble	☺	☺	☹	☺	☺
Siontis	☺	☹	☺	☺	☺
Sun	☹	☹	☹	☺	☹
Tzoulaki	☺	☹	☺	☺	☺
Yoshizawa	☺	☺	☹	☺	☺

Legend: ☺ = low risk; ☹ = high risk; **?** = unclear risk

Results of individual studies

The results for all the studies included are summarised qualitatively in Table3.

Table3–Summary of results

Authors	Year of publication	N° of studies included
Abbasi et al.	2012	16
Barber et al.	2014	12
Beswick et al.	2008	110
Collins et al.	2011	39
Cortes-Bergoderi	2012	5
Damen et al	2016	212
Echouffo-Tcheugui et al. 2015	2015	13
Echouffo-Tcheugui et al. 2013	2013	11
Fowkes	2008	20
Hu et al.	2016	12
Noble et al.	2011	43
Siontis et al.	2012	20
Sun	2017	26
Tzoulaki et al.	2009	79
Yoshizawa et al.	2016	18

Intervention	Comparison	Performance (AUC, C STAT., D-STAT)	Validation	Outcomes: incidence, prevalence, mortality	Quality assessment score
25 Type2Diabetes Risk Prediction Models	Prospective cohort study, with a case cohort study	C-statistics: 0.74-0.92	External validation cohort	Type 2 diabetes-morbidity (incidence)	6
18 Type2Diabetes Risk Prediction Models	Seven risk tools were validated using an external	AUROC: 0.69	Internal and external validation	Pre diabetes-morbidity (incidence)	6
110 risk scoring methods with potential for use in primary prevention	internal?	AUROC min 0.57, max 0.88	Internal and convergent validation	CHD, MI/suddenischaemicdeathCHD death, CVD, stroke, CHD, CVD deathMI or CHDdeath CHD death	11
Use of 47 different risk predictors	Ten studies randomly split the cohort into	NA	Internal and external validation	Diabetes-morbidity (incidence and prevalence)	6
Framingham Novel RPM based on local cohort WHO/ISH cardiovascular risk score	Framingham *not for Chagas	C-stat MAX 0.81 (95% CI, 0.72–0.90);	N\A (chagas only)	General mortality, CVD-mortality and morbidity, (incidence), and CHD-mortality and morbidity (incidence)	6
363 different models	internal?	For143 (39%) models, discrimination	80 of the 363 developed models (22%)	Fatal or non-fatal CHD Fatal or non-fatal: CHD, CVD, myocardial infarction, and stroke	7
22 RISK SCORES: FHS risk score Health ABC Score	internal?	AUC from 0.71 to 0.87	HL $\chi 2$ from 2.98 to 9.45	Heart Failure incidence	6
Johns Hopkins Framingham score Women's Health Study (WHS)	internal?	MAX 0.803, MIN 0.707	External? C-statistic in validation	Hypertension morbidity (incidence)	9
Baseline ABI measurements	Framingham risk score	N\A		General mortality, CVD-mortality, and CHD-morbidity (incidence)	6
12 Type2Diabetes Risk Prediction Models	to be seen	AUROC: 0.66-0.91	Internal and external validation	Diabetes-morbidity (incidence)	8
94 Type2Diabetes Risk Prediction Models		AUROC: 0.74-0.85	Internal and external validation	Diabetes-morbidity (incidence)	5
Framingham ASSIGN score SCORE score	internal?	AUC MIN 0.55 MAX 0.83	N\A	CVD mortality, CHD incidence, cerebrovascular disease incidence, CABG or PTCA	9
Anthropometric indices risk prediction ARIC/CHC risk score biomarker-based risk-prediction model	Internal?	AUC = 0.767, 95% CI(0.742, 0.792)	Variable	Hypertension morbidity (incidence)	9
Application of FRS plus a candidate additional predictor	Framingham risk score	AUC 0.77	n.p.	General mortality, CVD-mortality and morbidity (incidence), and CHD-mortality and morbidity (incidence)	5
Non-blood-based risk model for Type2 Diabetes	None of the included studies met all criteria on the	AUC: 0.71-0.79	Internal and external validation	Diabetes-morbidity (incidence)	8

Studies on Diabetes

Abbasi et al (18) (AMSTAR 6/11) focused on 16 prospective cohort studies, in order to validate 25 risk-predictive models for type 2 diabetes (T2DM), by means of an external validation cohort. The sample was composed by 38,379 people aged 20-70 with no diabetes at baseline. Incidence of type 2 diabetes was evaluated as outcome. All the studies included reported a C-statistic, ranging from 0.74 to 0.84 for risk at 7.5years, indicating a good discriminatory capability. The risk models had an estimate of calibration, the Hosmer-Lemeshow test, generally indicating good calibration.

Barber et al (19) (AMSTAR 6/11) assessed the applicability of 18 risk assessment tools in individuals with pre-diabetes. Their systematic review included 12 studies, with sample sizes ranging from 1,351 to 7,092. Incidence of pre-diabetes, defined according to American Diabetes Association (ADA) criteria, was considered as primary outcome.
Validation (either internal or external) of the risk scores was achieved by evaluating both discrimination and calibration. The internal C-statistic ranged from 0.66 to 0.75. Calibration was stated by Hosmer-Lemeshow goodness-of-fit test value, reported by only two studies and with discordant results.

Collins et al (20) (AMSTAR 6/11) evaluated risk prediction models for type 2 diabetes, including 39 studies comparing 47 different risk tools. The studies had a median sample size of 2,562 people, with an interquartile range from 1,426 to 4,965. No quantitative information was available on discrimination and calibration.

Hu et al (21) (AMSTAR 5/11) evaluated the effectiveness of risk-predictive models for type 2 diabetes in the Asian population. Their systematic review included 43 studies examining 12 risk-predictive models, derived from population samples ranging from 2,677 to 73,961.
Discrimination was evaluated by the area under the curve (AUC): this showed a high variability of results, with AUC ranging from 0.66 to 0.91.

Noble et al (22) (AMSTAR 5/11) conducted a systematic review assessing94 risk models for type 2 diabetes.
The 43 evaluated perspective cohort studies had a sample size ranging from 399 to 2.54 million people and were focused on incidence of diabetes as primary outcome. Some of the risk models had been externally validated on a different population. The C-statistic index showed high variability, fluctuating between not acceptable (0.60) and good quality (0.91) scores. The same results applied for calibration indicators.

Yoshizawa et al (23) (AMSTAR 8/11) focused on evaluating the predictive ability of a non blood-based risk prediction model for incidence ofT2DM. The 18 eligible studies included an overall number of 184,011 participants aged 42.4 to 68.4 years. Discrimination, evaluated by the area under the curve, was adequate to good (0.72-0.81).

Studies on CVDs

Cortes-Bergodieri et al (24) (AMSTAR 6/11) assessed the validity of Risk Prediction Models in Latin America and in US people of Hispanic descent. Their review included five cohort studies, comparing the Framingham score with three risk models for CVD and one for Chagas Disease, and investigating incidence and mortality as outcomes. Risk score calibration measured by C-statistic index was good, ranging from 0.69 to 0.80. While they openly admit that the Framingham score needs to be recalibrated for Latin American populations, they also recognise that evidence regarding CVD risk

models is “modest at best”. Indeed, all the studies included showed a ratio of predicted/observed that was not significant.

Tzoulaki et al (25) (AMSTAR: 5/11) focused on 79 studies on Framingham-based improving models, derived from populations of less than 1,000 to over 10,000 subjects from the USA and the UK. Incidence and mortality for coronary heart disease were measured as outcome.
The discrimination ability of the examined scores, evaluated by AUC, was variable from not acceptable to good: the FRS alone model showed an area under the curve between 0.50 and 0.83, whereas FRS with additional predictors ranged from .57 to 0.84.

Fowkes et al (26) (AMSTAR 6/11) evaluated the ankle-brachial index (ABI) as predictor of cardiovascular events and mortality, compared to the Framingham Risk Score. They included 20 prospective studies concerning general populations from the EU and the US, with sample sizes ranging from 554 to 14,109.
The combination of ABI and FRS risk prediction scores had a higher discriminating power compared to FRS alone (0.655 versus 0.646 among men, and 0.658 versus 0.605 among women).
Incidence of CVDs was assessed, as primary outcome, using adjusted hazard ratio estimates.
The study results showed that ankle-brachial index measurement can be used in addition to FRS in order to improve its predictive power, and the Authors suggested that a combined tool could be useful.

Siontis et al (27) (AMSTAR 9/11) performed a comparison of eight risk prediction models for cardiovascular disease. Their review included 20 prospective and retrospective studies, with sample sizes ranging from 403 to 1,072,800. The main outcomes considered were CVD mortality and CVD-related incidence. Concerning probability for prediction of outcome, there was significant variability among the studies, ranging from poor (0.55) to good (0.85).

Damen et al (28) (AMSTAR: 7/11) conducted a systematic review examining 212 studies, describing the development of 363 different prediction models for CVD and CHD. Sample size was extremely variable, ranging between 51 and 1,189,845 people, mainly from Europe, Canada and the USA.
Measures of predictive performance were reported only in 53% of the studies included, with discriminatory ability from 0.61 to 1.00.
In addition, an external validation was performedon136 articles and this most often concerned four models: Framingham, SCORE, QRISK and ATPIII.
The median discriminative ability was always acceptable (0.70-0.79), except for the ATPIII score (C statistic index: 0.66). Calibration, estimated as observed: expected ratio, ranged between 0.59 of Framingham-Wilson and 0.94 of QRISK.

Beswick et al(29) (AMSTAR 11/11) included 30 articles evaluating several risk prediction methods for coronary heart disease (CHD) and cardiovascular disease (CVD): 16 studies using convergent validation of Framingham-Anderson based methods and 21 comparisons using different risk scoring methods. The enrolled samples comprised 4,540 to over 205,000 people, aged 5 to 70 years, from US, Australia, Europe and India. Incidence and mortality for coronary heart disease and cardiovascular disease were estimated as primary outcomes.
Only the most recent updates to the Sheffield tables and the Joint British charts showed acceptable sensitivity and specificity compared with the Framingham-Anderson model.
In addition, Beswick et al. performed a second systematic review on external validation of Framingham-based risk scoring methods, based on 62 longitudinal or cross-sectional studies conducted on 112 different populations.
The articles included reported extreme variability in terms of discriminatory ability, with areas under the curve ranging from not acceptable (0.58) to good (0.85), although with better results in women than in men and in people with more recent baseline examinations.

Concerning calibration, the predicted:observed ratios ranged from an under-prediction of 0.43 to an over-prediction of 2.87. Generally speaking, under-prediction was greater in people at higher risk, such as subjects with a family history of premature cardiovascular disease, and lower in people at lower risk.

Echouffo-Tcheugui et al (30) (AMSTAR: 6/11) focused on 13 studies evaluating 28 heart failure risk prediction models. They were based on a US and European cohort of 725 to 359,947, aged over 18. Assessed scores had acceptable-to-good discriminatory ability, with C-statistics ranging between 0.71 and 0.87.
Calibration, when reported, was generally acceptable. Only two models were externally validated, showing modest-to-acceptable discrimination, with C-statistics from 0.61 to 0.79.

Studies on Hypertension

Sun et al (31) (AMSTAR 9/11) included 26 cohort studies on hypertension assessing 48 risk models, including both traditional risk factors -such as body mass index (BMI), age, smoking, blood pressure (BP) level and parental history of hypertension- with biochemical parameters and genetic factors. Evaluated articles were developed from samples mainly drawn from populations in the US, Eastern Asia and the UK and they included a population aged over 20, with a sample size ranging from 443 to 17.471. All the studies included reported a C-statistic index ranging from 0.74 to 0.79, indicating a good discriminatory capability. Furthermore, most of the studies had an estimate of calibration, by Hosmer-Lemeshow test, that never gave significant results.

Echouffo-Tchegui et al 2013 (32) (AMSTAR 9/11) assessed 11 prospective cohort studies evaluating 15 different risk models for hypertension, with population samples ranging from 1,135 to 11,407 derived from US and Eastern Asian populations. Incidence of hypertension was considered as primary outcome. C-statistic ranged from 0.80 to 0.70, indicating good performance and discrimination. Ten models also estimated calibration, using the Hosmer-Lemeshow test, and this generally reported good calibration.

DISCUSSION AND CONCLUSIONS

Strengths and limitations
To the best of our knowledge, this is the most comprehensive umbrella systematic review on risk prediction models for NCDs, such as diabetes and cardiovascular diseases, with high incidence, prevalence and mortality worldwide. In fact, no other umbrella systematic reviews are available on this issue. The only umbrella systematic review found (33) was focused solely on hypertension. Topics were selected according to these criteria:

1. Epidemiological relevance in terms of incidence and prevalence;
2. The significant link between diabetes, CVD and hypertension in terms of pathogenic pathway and clinical presentation.

This is also the reason why cancer was not considered among the inclusion criteria.
Many studies had been conducted on NCDs, in particular during the last decade, and the Authors have therefore chosen to use an umbrella methodology for systematic reviews and meta-analyses. Unfortunately, available studies, although apparently being of medium-high quality according to AMSTAR score (mean 7.14 out of 11, ranging from 5 to 11), were based on primary studies of debatable quality, with a large proportion of them lacking discrimination and calibration assessments. Again, publication bias and reporting bias could not be assessed, mostly because of the scarcity of evidence, which may have led to overestimation of effective sizes.

Likewise, due to the significant heterogeneity of study designs, the risk models involved and the outcomes reported, it was not possible to conduct a meta-analysis. The results were therefore reported narratively.

Summary of evidence

Developing a good predictive score in order to enable early identification of diabetes, hypertension and CVDs is a major public health concern, not just in high-income countries, because of the extremely high incidence and prevalence rates of these diseases and their upward trend worldwide, which results in massive consumption of social, health and economic resources.

Despite the amount of evidence on the issue, the average quality of the primary studies included is still poor: no available external validation or model calibration, no standardised study design and optimism bias are just some examples. Furthermore, a number of searches showed that older, and more limited, RPMs had better performance than newer, and more complex, models (27).

The majority of studies, and in particular those predicting diabetes, reported comparisons that were often achieved among scores that were very similar in prediction model tools, such as those differing for a very small number of items (sometimes only one) and focusing on the same population. Probably for this reason, no RPMs on diabetes and hypertension seemed to excel: no significant difference was found in the majority of studies on these two diseases.

Conversely, some promising differences among prediction tools were highlighted for CVDs and CHDs. The new RPMs investigated generally used Framingham scores as the main comparison tool. According to Fowkes et al. (26), the ankle-brachial index, in association with the Framingham tool, improved performance results, albeit only slightly.

In addition, QRISK scores provided some evidence of a certain superiority compared to Framingham, in particular in calibration and in discrimination performance (28).

However, it should be pointed out that Framingham-based methods underestimated risk in diabetics, socioeconomically deprived populations, and patients with a strong family history of premature cardiovascular disease (29).

Because of all the limitations described in the available studies and because no clear superiority of one predictive model was identified, it seems legitimate to question whether investing in new risk models is still a good practice, or if it would be a better approach to focus such efforts on external validation of the existing tools.

Conclusions and future perspectives

The wide range of available studies testing risk prediction models on CVDs, CHDs, hypertension and diabetes, which compare almost overlapping tools that often differ only for single entries, does not really contribute to increasing our knowledge of the issue and, rather, merely increases uncertainty.

Some more precise evidence is available only for cardiovascular disease prediction: the Framingham score, alone or in combination with the ankle-brachial index, and QRISK score can be confirmed as the gold standard.

Further efforts should not be concentrated on creating new scores, but rather on performing external validation of the existing ones. Promising future possibilities on the issue could then involve testing risk scores on wider samples and on certain target populations, with specific exposure risks and for which no robust scientific evidence is currently available, such as workers. These individuals could definitely benefit from early detection of chronic disease, since this is often worsened by occupational exposure and is a cause of disability and absence from work. Benefits could be further improved by supplementing the models with information on lifestyle, personal habits and family history (23).

FUNDING, CONFLICT OF INTEREST AND AUTHORS' AFFILIATION

The authors confirm that no sources of funding were obtained.

The authors declared no conflict of interest.

REFERENCES

1 Health statistics and information systems. Disease burden and mortality estimates. CAUSE-SPECIFIC MORTALITY, 2000–2016.Available at: http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html. Last access: 17 January 2019.

2 Mills KT, Bundy JD, Kelly TN et al. Global Disparities of Hypertension Prevalence and Control: A Systematic Analysis of Population-Based Studies From 90 Countries. *Circulation*. 2016 Aug 9;134(6):441-50.doi: 10.1161/CIRCULATIONAHA.115.018912.

3 Global Health Observatory (GHO) data. Raised blood pressure. Situation and trends, Available at: http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence_text/en/. Last access: 18 June 2018.

4 WHO - Health statistics and information systems. DISEASE BURDEN, 2000–2016. Available at: https://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html. Last access: 17 January 2019.

5 Shaw LJ, Goyal A, Mehta C, et al. 10-Year Resource Utilization and Costsfor Cardiovascular Care. *J Am CollCardiol*. 2018 Mar 13;71(10):1078-1089.doi: 10.1016/j.jacc.2017.12.064.

6 American Diabetes Association. Economic Costs of Diabetes in the U.S. in 2017. *Diabetes Care*. 2018 May;41(5):917-928. doi: 10.2337/dci18-0007. Epub 2018 Mar 22.

7 High Blood Pressure. Blood Pressure: Make Control Your Goal Infographic, Available at: https://www.cdc.gov/bloodpressure/infographic.htm. Last access: 18 June 2018.

8 Mehta S, Ghosh S, Sander S,et al. Differences in All-Cause Health Care Utilization and Costs in a Type 2 Diabetes Mellitus Population with and Without a History of Cardiovascular Disease. *J Manag Care Spec Pharm*. 2018 Mar;24(3):280-290.doi: 10.18553/jmcp.2018.24.3.280.

9 Kung HC, Xu J. Hypertension-related Mortality in the United States, 2000-2013. *NCHS Data Brief*. 2015 Mar;(193):1-8.

10 Gupta R, Gupta VP, Prakash H, et al. 25-Year trends in Hypertension prevalence, awareness, treatment, and control in an Indian urban population: Jaipur Heart Watch. (2018) *Indian Heart Journal*, 70 (6), pp. 802-807.doi: 10.1016/j.ihj.2017.11.011.

11 Bhatnagar P, Wickramasinghe K, Wilkins E, et al. Trends in the epidemiology of cardiovascular disease in the UK. *Heart*. 2016 Dec 15;102(24):1945-1952. doi: 10.1136/heartjnl-2016-309573.Epub 22 August 2016.

12 European Heart Network. European Cardiovascular Disease Statistics 2017, Available at: http://www.ehnheart.org/CvD-statistics/CvD-statistics-2017.html. Last access: 18 June 2018.

13 Petrie JR, Guzik TJ, Touyz RM. Diabetes, Hypertension, and Cardiovascular Disease: Clinical Insights and Vascular Mechanisms. *Can J Cardiol*. 2018 May;34(5):575-584.doi:10.1016/j.cjca.2017.12.005.Epub11 December 2017.

14 Moher D, Liberati A, Tetzlaff J, et al; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009 Jul 21;6(7):e1000097.doi: 10.1371/journal.pmed.1000097.Epub21 July 2009.

15 Sharif MO, Janjua-Sharif FN, Ali H, et al. Systematic reviews explained: AMSTAR-how to tell the good from the bad and the ugly. *Oral Health Dent Manag*. 2013 Mar;12(1):9-16.

16 The Joanna Briggs Institute Reviewers' Manual: 2014 edition / Supplement Methodology for JBI Umbrella Reviews. Available at: http://joannabriggs.org/assets/docs/sumari/ReviewersManual-Methodology-JBI_Umbrella%20Reviews-2014.pdf. Last access: 17 January 2019.

17 Whiting P, Savović J, Higgins JP, et al; ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J ClinEpidemiol*.2016 Jan;69:225-34.doi:10.1016/j.jclinepi.2015.06.005. Epub16 June 2015.

18 Abbasi A, Peelen LM, Corpeleijn E, et al. Prediction models for risk of developing type 2 Diabetes: systematic literature search and independent external validation study. *BMJ*. 2012 Sep 18;345:e5900.doi: 10.1136/bmj.e5900.

19 Barber SR, Davies MJ, Khunti K, et al. Risk assessment tools for detecting those with pre-Diabetes: a systematic review. *Diabetes Res ClinPract*. 2014 Jul;105(1):1-13. doi: 10.1016/j.diabres.2014.03.007.Epub18 March 2014.

20 Collins GS, Mallett S, Omar O, et al. Developing risk prediction models for type 2 Diabetes: a systematic review of methodology and reporting. *BMC Med*. 2011 Sep 8;9:103. doi: 10.1016/j.diabres.2014.03.007.

21 Hu PL, Koh YL, Tan NC. The utility of Diabetes risk score items as predictors of incident type 2 Diabetes in Asian populations: An evidence-based review. *Diabetes Res ClinPract*. 2016 Dec;122:179-189. doi: 10.1016/j.diabres.2016.10.019.Epub2 November 2016.

22 Noble D, Mathur R, Dent T, et al. Risk models and scores for type 2 Diabetes: systematic review. *BMJ*. 2011 Nov 28;343:d7163.doi: 10.1136/bmj.d7163.

23 Yoshizawa S, Kodama S, Fujihara K, et al. Utility of nonblood-based risk assessment for predicting type 2 Diabetes mellitus: A meta-analysis. *Prev Med*. 2016 Oct;91:180-187.doi: 10.1016/j.ypmed.2016.07.026.Epub26 July 2016.

24 Cortes-Bergoderi M, Thomas RJ, Albuquerque FN, et al. Validity of cardiovascular risk prediction models in Latin America and among Hispanics in the United States of America: a systematic review. *Rev PanamSaludPublica*. 2012 Aug;32(2):131-9.

25 Tzoulaki I, Liberopoulos G, Ioannidis JP. Assessment of claims of improved prediction beyond the Framingham risk score. *JAMA*. 2009 Dec 2;302(21):2345-52.doi: 10.1001/jama.2009.1757.

26 Fowkes FG, Murray GD, Butcher I, et al. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. *JAMA*. 2008 Jul 9;300(2):197-208.doi: 10.1001/jama.300.2.197.

27 Siontis GC, Tzoulaki I, Siontis KCet al. Comparisons of established risk prediction models for cardiovascular disease: systematic review. *BMJ*. 2012 May 24;344:e3318.doi: 10.1136/bmj.e3318.

28 Damen JA, Hooft L, Schuit E, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ*. 2016 May 16;353:i2416.doi: 10.1136/bmj.i2416.

29 Beswick AD, Brindle P, Fahey T, Ebrahim S. A Systematic Review of Risk Scoring Methods and Clinical Decision Aids Used in the Primary Prevention of Coronary Heart Disease (Supplement) [Internet]. *London: Royal College of General Practitioners (UK); 2008 May. National Institute for Health and Clinical Excellence: Guidance.*

30 Echouffo-Tcheugui JB, Greene SJ, Papadimitriou L, et al. Population risk prediction models for incident heart failure: a systematic review. *Circ Heart Fail*. 2015 May;8(3):438-47.doi:10.1161/CIRCHEARTFAILURE.114.001896.Epub3 March 2015.

31 Sun D, Liu J, Xiao L, et al. Recent development of risk-prediction models for incident Hypertension: An updated systematic review. *PLoS One*. 2017 Oct 30;12(10):e0187240. doi: 10.1371/journal.pone.0187240. eCollection 2017.

32 Echouffo-Tcheugui JB, Batty GD, Kivimäki M, et al. Risk models to predict Hypertension: a systematic review. *PLoS One*. 2013 Jul 5;8(7):e67370doi: 10.1371/journal.pone.0067370. Print 2013.

33 Pocock SJ, McCormack V, Gueyffier F, et al. A score for predicting risk of death from cardiovascular disease in adults with raised blood pressure, based on individual patient data from randomised controlled trials. *BMJ*. 2001 Jul 14;323(7304):75-81.

34 Karmali KN, Persell SD, Perel P, et al. Risk scoring for the primary prevention of cardiovascular disease. *Cochrane Database Syst Rev.* 2017 Mar 14;3:CD006887.

Figures Caption

Figure 1 – Study workflow

Figure 2 – Quality assessment scores

Figure 3 – Graphical representation of risk of bias according to the ROBIS tool

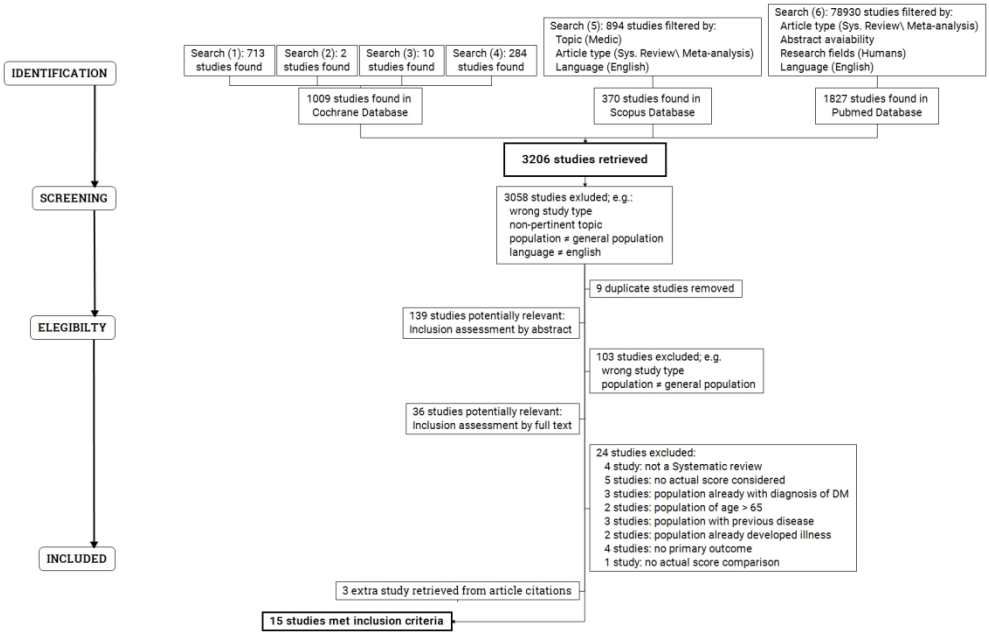

Figure 1 - Study workflow

154x99mm (300 x 300 DPI)

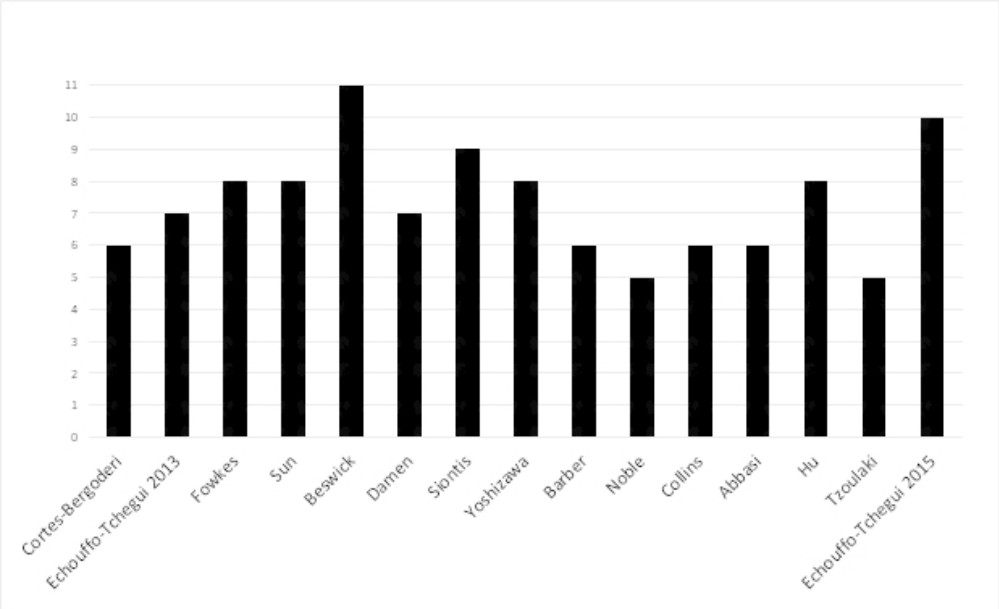

Figure 2 - Quality assessment scores

52x32mm (300 x 300 DPI)

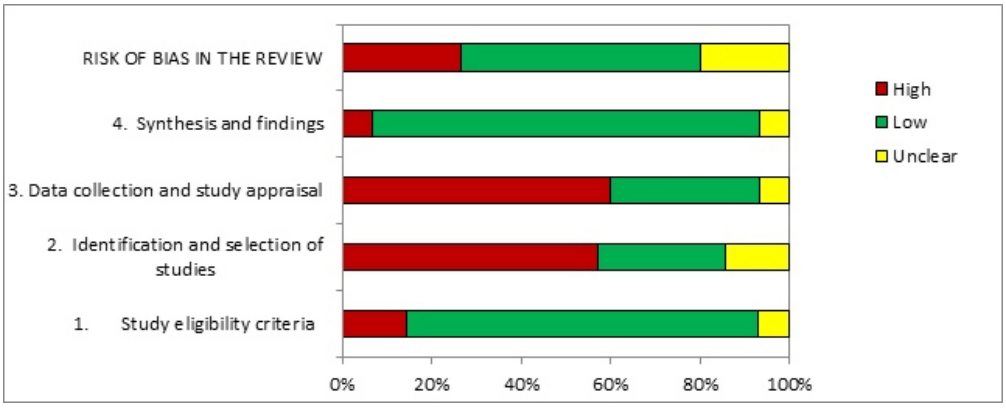

Figure 3 - Graphical representation of risk of bias according to the ROBIS tool

54x21mm (300 x 300 DPI)

Section/topic	#	Checklist item	Reported on page #
TITLE			
	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary			
	2	Provide a structured summary including, as applicable:	
		background;	1
		objectives;	1
		data sources;	1
		study eligibility criteria,	1
		participants, and	1
		interventions;	1
		study appraisal and synthesis methods;	
		results;	1
		limitations;	
		conclusions and implications of key findings;	2
		systematic review registration number.	2
INTRODUCTION			
Rationale			
	3	Describe the rationale for the review in the context of what is already known.	3
Objectives			
	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration			
	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria			
	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources			
	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search			
	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5
Study selection			
	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	4
Data collection process			
	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5
Data items			
	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4
Risk of bias in individual studies			
	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5
Summary measures			
	13	State the principal summary measures (e.g., risk ratio, difference in means).	NA
Synthesis of results			
	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis.	7
	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	NA
Additional analyses			
	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection			
	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	3
Study characteristics			
	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8
Risk of bias within studies			
	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	7
Results of individual studies			
	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8
Synthesis of results			
	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies			
	22	Present results of any assessment of risk of bias across studies (see Item 15).	7
Additional analysis			
	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence			

	24 Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	13
Limitations		
	25 Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	14
Conclusions		
	26 Provide a general interpretation of the results in the context of other evidence, and implications for future research.	14
FUNDING		
Funding		
	27 Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	15

BMJ Open

Can risk be predicted? An Umbrella Systematic Review of current Risk Prediction Models for Cardiovascular Diseases, Diabetes and Hypertension.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-030234.R1
Article Type:	Original research
Date Submitted by the Author:	08-Nov-2019
Complete List of Authors:	Lucaroni, Francesca; University of Rome Tor Vergata, Biomedicine and Prevention Cicciarella Modica, Domenico; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention Macino, Mattia; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention Palombi, Leonardo; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, biomedicina e prevenzione Abbondanzieri, Alessio; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention Agosti, Giulia; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention Biondi, Giorgia; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention Morciano, Laura; University of Rome Tor Vergata, Biomedicine and Prevention Vinci, Antonio; Università degli Studi di Roma Tor Vergata Facoltà di Medicina e Chirurgia, Biomedicine and Prevention
Primary Subject Heading:	Public health
Secondary Subject Heading:	Cardiovascular medicine, Epidemiology, Diabetes and endocrinology
Keywords:	Cardiovascular diseases, Hypertension < CARDIOLOGY, Diabetes, Risk Prediction Models

Can risk be predicted? An Umbrella Systematic Review of current Risk Prediction Models for Cardiovascular Diseases, Diabetes and Hypertension.

Prof. Leonardo Palombi, University of Rome Tor Vergata Department of Biomedicine and Prevention Rome, Lazio, IT Via Montpellier, 1 –00133– Rome, Italy
e-mail: palombi@uniroma2.it
Tel +39 3358436196

Domenico Cicciarella Modica, Mattia Macino, Antonio Vinci, Alessio Abbondanzieri, Giulia Agosti, Giorgia Biondi:
Università degli Studi di Roma Tor Vergata, Facoltà di Medicina e Chirurgia, Department of Biomedicine and Prevention, Roma, Lazio, IT

Francesca Lucaroni, Laura Morciano:

University of Rome Tor Vergata Department of Biomedicine and Prevention Rome, Lazio, IT

Word count: 5,600

Keywords: *Cardiovascular diseases*; *Hypertension*; *Diabetes*; *Risk Prediction Models*;

Can risk be predicted? An Umbrella Systematic Review of current Risk Prediction Models for Cardiovascular Diseases, Diabetes and Hypertension.

Lucaroni F[2], Cicciarella Modica D[1], Macino M[1], Palombi L[2*], Abbondanzieri A[1], Agosti G[1], Biondi G[1], Morciano L[2], Vinci A[1].

[1]School of Hygiene and Preventive Medicine, University of Rome Tor Vergata – Rome, Italy
[2]Department of Biomedicine and Prevention, University of Rome Tor Vergata – Rome, Italy

*: Corresponding author

Abstract

Objective: to provide an overview of the currently available risk prediction models (RPMs) for Cardiovascular diseases (CVDs), Diabetes and Hypertension, and to compare their effectiveness in proper recognition of patients at risk of developing these diseases.

Design: Umbrella Systematic Review

Data Sources: PubMed, Scopus, Cochrane Library

Eligibility Criteria: Systematic Reviews or Meta-Analysis examining and comparing performances of Risk Prediction Models for Cardiovascular Diseases, Hypertension or Diabetes in healthy adult (18-65 years old) population, published in English language.

Data extraction and synthesis: data were extracted according to the following parameters: number of studies included, intervention (RPMs applied/assessed), comparison, performance, validation and outcomes. A narrative synthesis was performed. Data were reported according to the PRISMA guidelines.

Study Selection: 3612 studies were identified. After title/abstract screening and removal of duplicate articles, 37 studies met the eligibility criteria. After reading the full-text, 13 were deemed relevant for inclusion. Three further papers from the reference lists of these articles were then added.

Study Appraisal: the methodological quality of the included studies was assessed using the AMSTAR tool.

Risk of Bias in individual studies: Risk of Bias evaluation was carried out using the ROBIS tool.

Results: 16 studies met the inclusion criteria: 6 focused on Diabetes, 2 on Hypertension and 8 on CVDs.
Globally, prediction models for Diabetes and Hypertension showed no significant difference in effectiveness. Conversely, some promising differences among prediction tools were highlighted for CVDs. The ankle-brachial index, in association with the Framingham tool, and QRISK scores provided some evidence of a certain superiority compared to Framingham alone.

Limitations: due to the significant heterogeneity of the studies, it was not possible to perform a meta-analysis.
The electronic search was limited to studies in English and to three major international databases (MEDLINE/PubMed, Scopus and Cochrane Library), with additional works derived from the

reference list of other studies; grey literature with unpublished documents was not included in the search. Furthermore, no assessment of potential adverse effects of RPMs was carried out.

Conclusions: consistent evidence is available only for cardiovascular disease prediction: the Framingham score, alone or in combination with the ankle-brachial index, and the QRISK score can be confirmed as the gold standard.
Further efforts should not be concentrated on creating new scores, but rather on performing external validation of the existing ones, in particular on high risk groups.
Benefits could be further improved by supplementing existing models with information on lifestyle, personal habits, family and employment history, social network relationships, income and education.

PROSPERO SRRN: CRD42018088012

ARTICLE SUMMARY:

Strengths and Limitations

- This is the most comprehensive umbrella systematic review on RPMs for CVDs, Hypertension and Diabetes to date;
- Available studies, although apparently of medium-high quality, were based on primary studies of debatable quality, several of which lack discrimination and calibration assessments;
- Grey literature was not searched
- Heterogeneity was too high for meta-analysis; results are therefore reported narratively.

INTRODUCTION

Cardiovascular diseases (CVDs), hypertension and diabetes represent a major health concern, throughout the world and across income levels, as a silent epidemic responsible for millions of deaths every year.

CVDs, excluding hypertension are the leading cause of death worldwide; they–have a global prevalence rate of 6.6% and account for 17.6 million deaths per year (1).

Hypertension alone is the leading preventable cause of premature death worldwide (2), and causes 7.5 million deaths each year (3).
In 2010, almost one of every three adults (31.1%) had hypertension, although there was a significant gap in prevalence rates between high- and low-income countries: 28.5% (27.3%-29.7%) and 31.5% (30.2-32.9%) respectively worldwide (2).

Diabetes (types 1 and 2 combined) has a global prevalence rate of 5.4% and is responsible for 1.4 million deaths every year (3).
In addition to being a significant cause of mortality worldwide, CVDs, hypertension and diabetes are also a leading cause of disability. Together they account for almost 40% of disability-adjusted life years (DALYs): cardiovascular diseases alone are responsible for 32.3% of total DALYs, hypertension for 3.7% and diabetes for 2.4% (4).

The epidemiology of these three diseases helps explain the substantial economic impact they have on national health services: in the United States, cardiovascular disease-related direct costs amount to approximately $444 billion per year (5), whereas the costs of diabetes are estimated as $327 billion per year; hypertension has an annual estimated cost of $51 billion (6), most of which (nearly $48 billion) represents direct medical expenses (7).
In recent decades, hypertension and diabetes have shown an increasing trend in both prevalence and mortality rates. It has been estimated that the prevalence of diabetes will continue to grow: 1 in 5 to 1 in 3 adults will be affected by 2050. The same is true for mortality rate trends: in the last fifteen years, diabetes-related deaths increased by 1%, and these data are expected to climb dramatically over the coming decades (8).
The prevalence of hypertension and the associated mortality rates have increased significantly in bothmen and women (9), and by 2030, prevalence is projected to be 44% (10).
Conversely, prevalence and mortality rate trends for CVDs have significantly decreased (11), including in countries that had seen considerable increases until the beginning of the 21st century (12).

Cardiovascular diseases, hypertension and diabetes are strongly related to each other: diabetes is associated with an increased risk of CVDs, which is exaggerated by concomitant hypertension. These conditions also share the same pathogenic pathways, at both macroscopic and molecular level: oxidative stress, inflammation and fibrosis, cause microvascular and macrovascular complications in diabetes, and also lead to vascular remodelling and dysfunctions in hypertension (13).
Because of the high prevalence and mortality rates associated with CVDs, hypertension and diabetes, and their related direct and indirect costs, early identification of individuals at high risk for these diseases is crucial; it results in terms of obtaining significant savings in both global health outcomes and economic expenditures.
A number of prediction models focused on these three non-communicable diseases (NCDs) is available, but there is no consensus as to the gold standard tools best utilised in practice.
The aim of this study is to provide an overview of the currently available risk prediction models (RPMs) for CVDs, Diabetes and Hypertension and to compare their effectiveness at properly recognising vulnerable people, at risk of developing these NCDs.

METHODS

This umbrella systematic review was performed following a protocol designed *a priori*, and reported according to the PRISMA guidelines (14).
The PubMed, Scopus and Cochrane databases were searched electronically on, September 24th 2019, using combinations of the relevant medical subject heading (MeSH) terms, key words, and word variants for "risk prediction scores" and "cardiovascular disease", "diabetes" and "hypertension", as shown in Table 1. The search and selection criteria were restricted to the following: systematic reviews, with or without meta-analysis, as the type of study; general population aged 18-65 with no major illness; comparison of at least two risk prediction models; English language. For the PubMed database only two further filters were added: only articles on humans and with abstracts available were included. No restrictions were applied in terms of publication date in any of the databases.
Reference lists of relevant articles and reviews were hand-searched for additional reports.
The study protocol was registered on the PROSPERO (*International prospective register of systematic reviews*) database (Registration number: CRD42018088012).

Table1–Search strategy

Database	Search string(s)	Filters
Cochrane Library	risk (prediction OR assessment) (model OR score)	/
	diabetes risk prediction	
	cardiovascular risk prediction	
	hypertension risk prediction	
PubMed	(risk[Title/abstract] OR "risk assessment"[MeSH Terms]) AND predict*[Title/abstract] AND (model[Title/abstract] OR score[Title/abstract])	Systematic reviews
		Meta-analysis
		Abstract
		Humans
		English
Scopus	(risk AND prediction AND model AND (systematic AND review OR meta-analysis))	Review
		English
		Medic

Two different authors independently screened the article titles and abstracts in each database: MM and AA for Scopus, AV and DCM for PubMed/Medline, AV and AA for Cochrane Library.
Disagreements were discussed by the authors and resolved by consensus or by recourse to a third author (FL). Studies were then labelled for inclusion or exclusion.
Every article meeting the eligibility criteria -Systematic Reviews/ Meta-Analyses on risk prediction models on CVDs, Diabetes and Hypertension, evaluated by comparison with other RPMs, in adults with no relevant illness- was considered for subsequent qualitative synthesis; duplicate records were removed, as were articles that included the exclusion criteria: any study carried out with the sole purpose of developing a new RPM, or validating one, or to propose a diagnostic/prognostic tool, without any comparisons to other prediction models.
Studies were henceforth labelled for inclusion or exclusion.
The selection process described above is summarised by the flow diagram shown in Figure 1.

Data extraction

Four authors (AA, DCM, MM and AV) extracted the data, including the following variables: number of the included studies; intervention (risk prediction models); comparison; performance: Area Under the receiver-operating characteristic Curve (AUC), C-statistic, D-statistic); validation: internal, external, both or not provided; outcomes: incidence, prevalence, mortality.

Assessment of study quality

Quality assessment of individual studies was performed by applying the AMSTAR tool (15). According to their score, articles were classified into three groups: low (AMSTAR score <4), medium (AMSTAR score ≥ 4 and ≤ 7), and high quality (AMSTAR score ≥ 8).
Four authors (AA, DCM, MM and AV) independently assigned score. Disagreements were resolved by consensus or by discussion with a fifth author (FL). No reviews were excluded *ex-post* for quality reasons.
A qualitative (narrative) synthesis of the selected reviews was then performed; the guidelines for umbrella review from the Joanna Briggs Institute were applied (16).

Risk of bias in individual studies

Risk of bias in individual studies was independently evaluated by four authors (AA, DCM, MM and AV), using the ROBIS tool (17). Any disagreements were resolved by consensus or by discussion with a fifth author (FL).

Patient and Public Involvement

As the study design was a systematic review, neither patients nor the public were involved.

RESULTS

A total of 3612 studies were identified through a search of the electronic databases.
After title screening, 3463 studies were excluded because they did not meet the eligibility criteria and nine duplicate articles were removed. Of the-149 studies passing the first evaluation stage, 102 studies were excluded because the topic was not pertinent, sample characteristics were inadequate (i.e. some articles included non-healthy populations), the article type did not meet the inclusion criteria (i.e. some studies were not systematic reviews), or there was no comparison between models.
After reading the full text of the 37 remaining studies, 13 were deemed relevant for inclusion. Twenty-four studies were excluded because of a lack of comparison model, related but non-pertinent topic or wrong article type.
Three additional studies were included after consulting reference lists of relevant articles and reviews overall, 16 studies met the eligibility criteria and were included in qualitative synthesis.

Quality assessment of the studies

According to the AMSTAR tool, all of the studies were of medium-to-high quality, with a mean score of 7.14, (range 5-11).Specifically, 8 of 15 (53.4%) were of medium quality (AMSTAR ≥ 4 and ≤ 7), and 7 of 15 (46.6%) were of high quality (AMSTAR > 8).
The results of quality assessment have been summarised in graph form (Figure 2).

Risk of bias within studies

The results of the risk-of-bias evaluation are shown in Table 2 and Figure 3.
Globally, only nine articles of sixteen (56.25%) had a low risk of bias, according to the ROBIS tool. Two articles had a high risk of bias due to the eligibility criteria: search limitation on English language. Eight articles had a high risk of bias due to the identification and selection process: the most common source of bias was the search limitation to a single database. Nine articles had a high risk of bias due to data collection and study appraisal, in particular because of the lack of formal appraisal tools. Finally, one article had a high risk of bias due to the synthesis and identification

process, mainly due to significant heterogeneity of primary sources. Given the overall heterogeneity of the retrieved studies, it was impossible to conduct a meta-analysis.

Risk of bias across systematic reviews

We identified all of the ongoing SRs that met our inclusion criteria by searching the PROSPERO database, in order to assess publication bias.

Eight ongoing systematic reviews were found, including the present study, and none have been published to date.

Table2 - Risk of bias

Review	Phase 2				Phase 3
	1. Study eligibility criteria	2. Identification and selection of the studies	3. Data collection and study appraisal	4. Synthesis and findings	Risk of bias in the systematic reviews included
Abbasi	☺	☹	☹	☺	☹
Barber	☺	☹	☹	☺	☹
Beswick	☺	☺	☺	☺	☺
Collins	☺	☹	☹	☺	☺
Cortes-Bergoderi	☺	☹	☹	☺	☹
Damen 2016	☹	☹	☺	☺	☺
Damen 2019	☺	☺	☺	☺	☺
Echouffo-Tcheugui 2013	☺	☺	☹	☺	☺
Echouffo-Tcheugui 2015	☺	?	☺	☺	?
Fowkes	☺	☺	☹	?	?
Hu	?	?	?	☹	?
Noble	☺	☺	☹	☺	☺
Siontis	☺	☹	☺	☺	☺
Sun	☹	☹	☹	☺	☹
Tzoulaki	☺	☹	☺	☺	☺
Yoshizawa	☺	☺	☹	☺	☺

Legend: ☺ = low risk; ☹ = high risk; **?** = unclear risk

Results of individual studies

Overall, 16 studies met the inclusion criteria, 8 of them concerning Diabetes, 6 CVDs and 2 Hypertension.

The results for all the included studies are summarised qualitatively in Table3.

Table3–Summary of results. AUC, AUROC =Area Under the Curve/Area Under the Receiver-Operating characteristic Curve; C-STAT = C-statistic; D-STAT = D-statistic.

Authors	Year of publication
Abbasi et al.	2012
Barber et al.	2014
Beswick et al.	2008
Collins et al.	2011
Cortes-Bergoderi	2012
Damen et al	2016
Damen et al.	2019
Echouffo-Tcheugui et al. 2015	2015
Echouffo-Tcheugui et al. 2013	2013
Fowkes	2008
Hu et al.	2016
Noble et al.	2011
Siontis et al.	2012
Sun	2017
Tzoulaki et al.	2009
Yoshizawa et al.	2016

N° of studies included	Intervention	Comparison	Performance (AUC, C STAT., D-STAT)	Validation	Outcomes: incidence, prevalence, mortality
16	25 Type 2 Diabetes Risk Prediction Models	Prospective cohort study, with a case cohort study random subcohort	C-statistics: 0.74-0.92	External validation cohort	Type 2 diabetes-morbidity (incidence)
12	18 Type 2 Diabetes Risk Prediction Models	Seven risk tools were validated using an external dataset. Fourteen of	AUROC: 0.69	Internal and external validation cohort	Pre diabetes-morbidity (incidence)
110	110 risk scoring methods with potential for use in primary prevention (among them, 70 methods specifically	internal?	AUROC min 0.57, max 0.88	Internal and convergent validation	CHD, MI/sudden ischaemic deathCHD death, CVD, stroke, CHD, CVD death MI or CHDdeath CHD death CHD
39	Use of 47 different risk predictors	Ten studies randomly split the cohort into development and	NA	Internal and external validation cohort	Diabetes-morbidity (incidence and prevalence)
5	Framingham Novel RPM based on local cohort WHO/ISH cardiovascular risk score custom	Framingham *not for Chagas	C-stat MAX 0.81 (95% CI, 0.72–0.90); MIN 0.69	N\A (eChagas only)	General mortality, CVD-mortality and morbidity, (incidence), and CHD-mortality and morbidity (incidence)
212	363 different models	internal?	For143 (39%) models, discrimination	80 of the 363 developed models (22%)	Fatal or non-fatal CHD Fatal or non-fatal: CHD, CVD, myocardial infarction, and stroke
38	Framingham Adult Treatment Panel (ATP) III, the Framingham Wilson model and the pooled cohort equations	intternal	OE from 0.58 to 0.79; C-STAT from	External	Fatal or non-fatal MI, Fatal or non-fatal stroke, Fatal, Hard or non-fatal CHD, angina pectoris, TIA, ASCVD
13	22 RISK SCORES: FHS risk score Health ABC Score Kaiser Permanente	internal?	AUC from 0.71 to 0.87	HL $\chi 2$ from 2.98 to 9.45	Heart Failure incidence
11	Johns Hopkins Framingham score Women's Health Study (WHS) Inclusive model	internal?	MAX 0.803, MIN 0.707	External? C-statistic in validation studies (0.71 to	Hypertension morbidity (incidence)
20	Baseline ABI measurements	Framingham risk score	N\A		General mortality, CVD-mortality, and CHD-morbidity (incidence)
12	12 Type 2 Diabetes Risk Prediction Models	to be seen	AUROC: 0.66-0.91	Internal and external validation cohort	Diabetes-morbidity (incidence)
43	94 Type 2 Diabetes Risk Prediction Models		AUROC: 0.74-0.85	Internal and external validation cohort	Diabetes-morbidity (incidence)
20	Framingham ASSIGN score SCORE score PROCAM score	internal?	AUC MIN 0.55 MAX 0.83	N\A	CVD mortality, CHD incidence, cerebrovascular disease incidence, CABG or PTCA
26	Anthropometric indices risk prediction ARIC/CHC risk score biomarker-based risk-prediction model China risk prediction model 1	Internal?	AUC = 0.767, 95% CI(0.742, 0.792)	Variable	Hypertension morbidity (incidence)
79	Application of FRS plus a candidate additional predictor	Framingham risk score	AUC 0.77	n.p.	General mortality, CVD-mortality and morbidity (incidence), and CHD-mortality and morbidity (incidence)
18	Non-blood-based risk model for Type 2Diabetes	None of the included studies met all criteria on the QUADAS list	AUC: 0.71-0.79	Internal and external validation cohort	Diabetes-morbidity (incidence)

8	5	9	9	5	8	6	6	6	9	7	6	6	11	6	6	Quality assessment score

Studies on Diabetes

Abbasi et al (18) (AMSTAR 6/11) focused on 16 prospective cohort studies, in order to validate 25 risk-predictive models for type 2 diabetes (T2DM), by means of an external validation cohort. The sample was included 38,379 people aged 20-70 with no diabetes at the baseline. Incidence of type 2 diabetes was evaluated as outcome. All included studies reported a C-statistic, ranging from 0.74 to 0.84 for risk at 7.5 years, indicating a good discriminatory capability. The risk models had an estimate of calibration, the Hosmer-Lemeshow test, generally indicating good calibration.

Barber et al (19) (AMSTAR 6/11) assessed the applicability of 18 risk assessment tools in individuals with pre-diabetes. Their systematic review included 12 studies, with sample sizes ranging from 1,351 to 7,092. Incidence of pre-diabetes, defined according to American Diabetes Association (ADA) criteria, was considered as the primary outcome.
Validation (either internal or external) of the risk scores was achieved by evaluating both discrimination and calibration. The internal C-statistic ranged from 0.66 to 0.75. Calibration was described by Hosmer-Lemeshow goodness-of-fit test value, reported by only two studies and with discordant results.

Collins et al (20) (AMSTAR 6/11) evaluated risk prediction models for type 2 diabetes, including 39 studies comparing 47 different risk tools. The studies had a median sample size of 2,562 people, with an interquartile range from 1,426 to 4,965. No quantitative information was available on discrimination or calibration.

Hu et al (21) (AMSTAR 5/11) evaluated the effectiveness of risk-predictive models for type 2 diabetes in the Asian population. Their systematic review included 43 studies examining 12 risk-predictive models, derived from population samples ranging from 2,677 to 73,961.
Discrimination was evaluated by the area under the curve (AUC): this showed a high variability of results (AUC 0.66 to 0.91).

Noble et al (22) (AMSTAR 5/11) conducted a systematic review assessing 94 risk models for type 2 diabetes.
They evaluated 43 perspective cohort studies (sample size ranging from 399 to 2.54 million people) and incidence of diabetes was the primary outcome. Some of the risk models had been externally validated on a different population. The C-statistic index showed high variability, fluctuating between not acceptable (0.60) and good quality (0.91) scores. The same results applied for calibration indicators.

Yoshizawa et al (23) (AMSTAR 8/11) focused on evaluating the predictive ability of a non blood-based risk prediction model for incidence of T2DM. The 18 eligible studies included an overall number of 184,011 participants aged 42.4 to 68.4 years. Discrimination, evaluated by the area under the curve, was adequate to good (0.72-0.81).

Studies on CVDs

Cortes-Bergoderi et al (24) (AMSTAR 6/11) assessed the validity of Risk Prediction Models in Latin America and in US people of Hispanic descent. Their review included five cohort studies, comparing the Framingham score with three risk models for CVD and one for Chagas Disease, and investigating incidence and mortality as outcomes. Risk score calibration measured by C-statistic index was good, (0.69 to 0.80). While the Authors openly admit that the Framingham score needs to be recalibrated for Latin American populations, they also recognise that evidence regarding CVD risk models is "modest at best". Indeed, all of the included studies showed a ratio of predicted/observed that was not significant.

Tzoulaki et al (25) (AMSTAR: 5/11) focused on 79 studies on Framingham-based improving models, derived from populations of less than 1,000 to over 10,000 subjects from the USA and the UK. Incidence and mortality for coronary heart disease were measured as outcome.
The discrimination ability of the examined scores, evaluated by AUC, varied from not acceptable to good: the FRS alone model showed an area under the curve between 0.50 and 0.83, whereas FRS with additional predictors ranged from .57 to 0.84.

Fowkes et al (26) (AMSTAR 6/11) evaluated the ankle-brachial index (ABI) as predictor of cardiovascular events and mortality, compared to the Framingham Risk Score. They included 20 prospective studies involving general populations from the EU and the US, with sample sizes that ranged from 554 to 14,109.
The combination of ABI and FRS risk prediction scores had a higher discriminating power compared to FRS alone (0.655 versus 0.646 among men, and 0.658 versus 0.605 among women).
Incidence of CVDs was assessed, as primary outcome, using adjusted hazard ratio estimates.
The study results showed that ankle-brachial index measurement can be used in addition to FRS to improve its predictive power, and the Authors suggested that a combined tool could be useful.

Siontis et al (27) (AMSTAR 9/11) performed a comparison of eight risk prediction models for cardiovascular disease. Their review included 20 prospective and retrospective studies, with sample sizes ranging from 403 to 1,072,800. The main outcomes considered were CVD mortality and CVD-related incidence. The probability for prediction of outcome varied significantly among the studies, from poor (0.55) to good (0.85).

Damen et al (28) (AMSTAR: 7/11) conducted a systematic review examining 212 studies, that described the development of 363 different prediction models for CVD and CHD. Sample size was extremely variable, ranging between 51 and 1,189,845 people, mainly from Europe, Canada and the USA.
Measures of predictive performance were reported in 53% only of the studies, with discriminatory ability from 0.61 to 1.00.
In addition, an external validation test was performed on 136 articles and most often concerned four models: Framingham, SCORE, QRISK and ATPIII.
The median discriminative ability was always acceptable (0.70-0.79), except for the ATPIII score (C statistic index: 0.66). Calibration was estimated as observed: expected ratio, ranged from 0.59 of Framingham-Wilson to 0.94 of QRISK.

Beswick et al (29) (AMSTAR 11/11) included 30 articles that evaluated several risk prediction methods for coronary heart disease (CHD) and cardiovascular disease (CVD): 16 studies using convergent validation of Framingham-Anderson based methods and 21 comparisons used different risk scoring methods. The enrolled samples involved 4,540 to over 205,000 people, aged 5 to 70

years, from US, Australia, Europe and India. Incidence and mortality for coronary heart disease and cardiovascular disease were estimated as primary outcomes.
Only the most recent updates to the Sheffield tables and the Joint British charts showed acceptable sensitivity and specificity compared to the Framingham-Anderson model.
In addition, Beswick et al. performed a second systematic review of external validation of Framingham-based risk scoring methods, based on 62 longitudinal or cross-sectional studies conducted on 112 different populations.
The results indicated extreme variability in discriminatory ability, with areas under the curve ranging from not acceptable (0.58) to good (0.85), the results were better in women than in men or in people with more recent baseline examinations.
Concerning calibration, the predicted:observed ratios ranged from an under-prediction of 0.43 to an over-prediction of 2.87. Generally speaking, under-prediction was greater in people at higher risk, such as subjects with a family history of premature cardiovascular disease, and lower in people at lower risk.

Echouffo-Tcheugui et al (30) (AMSTAR: 6/11) focused on 13 studies that evaluated 28 heart failure risk prediction models. Studies were based on a US and European cohort of 725 to 359,947 subjects, over 18 years of age. Assessed scores had acceptable-to-good discriminatory ability, with C-statistics ranging between 0.71 and 0.87.
Calibration, when reported, was generally acceptable. Only two models were externally validated and showed modest-to-acceptable discrimination, with C-statistics from 0.61 to 0.79.

Damen et al (31) (AMSTAR:9/11) included 38 studies and compared the performance of the Framingham Adult Treatment Panel (ATP) III, the Framingham Wilson model and the pooled cohort equations (PCE) for fatal or non-fatal coronary heart disease (Framingham Wilson and ATP III) and hard atherosclerotic CVD (PCE). Results for men and women were compared separately. The authors performed meta-analyses of the included studies, calibration was assessed through the observed vs expected (OE) ratio and discriminative power, through the C-statistic, for 10-year risk predictions. The OE ratio results were very heterogeneous, ranging from 0.58 to 0.79. C-statistic values were highly variable as well (from 0.58 to 0.82). Most of the studies showed overprediction of the expected events, especially in high-risk groups. According to the authors, risk prediction models for CVDs and CHDs showed a similar performance.

Studies on Hypertension

Sun et al (32)(AMSTAR 9/11) included 26 cohort studies on hypertension that assessed 48 risk models and included both traditional risk factors - body mass index (BMI), age, smoking, blood pressure (BP) level and parental history of hypertension- with biochemical parameters and genetic factors. Evaluated articles were based on samples drawn mainly from populations in the US, Eastern Asia and the UK and included a population aged over 20, with a sample size that ranged from 443 to 17.471. All the studies included reported a C-statistic index ranging from 0.74 to 0.79, indicating a good discriminatory capability. Furthermore, calibration estimates of most studies by Hosmer-Lemeshow test showed no significant results.

Echouffo-Tchegui et al 2013(33) (AMSTAR 9/11) assessed 11 prospective cohort studies that evaluated 15 different risk models for hypertension in population samples from 1,135 to 11,407 subjects from US and Eastern Asian populations. Incidence of hypertension was considered as primary outcome. The C-statistic ranged from 0.80 to 0.70, indicating good performance and discrimination. Ten models also estimated calibration,—using the Hosmer-Lemeshow test, and generally reported good calibration.

DISCUSSION AND CONCLUSIONS

Summary of evidence

Developing a good predictive score to enable early identification of diabetes, hypertension and CVDs is a major public health concern across countries of all income levels because of the extremely high rates of incidence and prevalence of these diseases, their upward trend worldwide and their massive consumption of social, health and economic resources.

In spite of the amount of evidence on the issue, the average quality of the existing primary studies is poor: they lack external validation, model calibration and standardized study design, and suffer from optimism bias. In fact, a number of searches showed that older and more limited RPMs performed better than newer, more complex models (27).

The majority of studies, in particular those predicting diabetes, reported comparisons that were often achieved with scores that were very similar in prediction model tools, including those differing by a very small number of items (sometimes only one) and those focused on the same population. It is therefore not surprising that no RPMs on diabetes and hypertension have seemed to excel: no significant difference was found in the majority of studies on these two diseases.

Conversely, some promising differences among prediction tools were highlighted for CVDs and CHDs. The new RPMs investigated generally used Framingham scores as the main comparison tool. According to Fowkes et al. (26), the ankle-brachial index, in association with the Framingham tool, improved performance results, albeit only slightly. In addition, QRISK scores provided some evidence of superiority compared to Framingham, in particular in the areas of calibration and discrimination performance (28). However, it should be pointed out that Framingham-based methods underestimated risk in diabetics, socioeconomically deprived populations, and in patients with a strong family history of premature cardiovascular disease (29). Because of the limitations described in the available studies, and because no predictive model was clearly identified as superior, it seems legitimate to question whether investing in new risk models is still a good practice, or if it would be a better approach to focus our efforts on external validation of existing tools.

Strengths and limitations

To the best of our knowledge, this is the most comprehensive umbrella systematic review on risk prediction models for NCDs, such as Diabetes and Cardiovascular Diseases, with high incidence, prevalence and mortality worldwide. In fact, no other umbrella systematic reviews are available on this issue. The only umbrella systematic review found (34) was focused solely on Hypertension. For this review topics were selected according to the following criteria:

1. Epidemiological relevance in terms of incidence and prevalence;
2. The significant link between Diabetes, CVD and Hypertension in terms of pathogenic pathway and clinical presentation.

This is also the reason why cancer was not considered among the inclusion criteria.

Many studies had been conducted on NCDs, in particular during the last decade, and the Authors have therefore chosen to use an umbrella methodology for systematic reviews and meta-analyses. Unfortunately, available studies, although reported to be of medium-high quality according to AMSTAR score (mean 8.07 out of 11, ranging from 5 to 11), were based on primary studies of debatable quality, with a large proportion of them lacking discrimination and calibration assessments. Due to the significant heterogeneity of study designs, the risk models involved and the outcomes reported, it was not possible to conduct a meta-analysis. The results were therefore reported narratively.

Our electronic search was limited to studies in English and to three major international databases (MEDLINE/PubMed, Scopus and Cochrane Library), with additional works derived from the reference list of studies, and did not include grey literature with unpublished documents. However, a

publication bias was estimated by the quantification of ongoing and non-completed systematic reviews in the PROSPERO database.
No assessment of potential adverse effects of RPMs has been carried out, with a potential risk of bias.

General interpretation of results

The increasing global growth in prevalence of chronic diseases, as a direct consequence of epidemiologic transition, has led to broader use of predictive tools as a major aid for health workers. Indeed, these instruments can be very important and should be regularly implemented in medical settings to support the activity of general practitioners and public health authorities involved in monitoring and evaluation of patients. Specific benefits of RPMs could emerge in prevention and health promotion for specific populations -such as workers and students- and social settings.
It must be pointed out that the studies evaluated in this systematic review, albeit of medium-high quality, are not primary studies and, therefore could be affected by significant bias. Furthermore, there is no evidence in the scientific literature for the evaluation of the effectiveness of RPMs on long-term patient outcomes (35). Therefore, the results from this study should be carefully applied by health workers, in order to minimize the risk of over- or under- treatment. (36).
Scientific literature in the past 30 years has produced an abundance of evidence on other powerful health determinants, such as social relationships networks, stress, unemployment, education and income (37-40), however none of these variables have been included in all the available predictive tools. Moreover, very few instruments considered lifestyle variables like smoking, alcohol, physical activity, and drug use or addiction. A strictly biological perspective should be considered as a serious limitation in terms of forecasting and predicting the development of CVDs, CHDs, diabetes and hypertension. A new generation of predictive tools, conceptually developed around biological and non-biological determinants, could consistently ameliorate the assessment of risk and the detection of risk stratification groups.

Conclusions and future perspectives

The wide range of available studies that have tested risk prediction models for CVDs, CHDs, hypertension and diabetes compare almost overlapping tools (which often differ by only a single entry), does not really increase our knowledge of the issue; rather it merely increases uncertainty.
More precise evidence is available only for cardiovascular disease prediction: the Framingham score, alone or in combination with the ankle-brachial index, and QRISK score can be confirmed as the gold standard.
Further efforts should not be concentrated on creating new scores, but rather on performing external validation of the existing ones. Promising future possibilities could then involve testing risk scores on wider samples and on certain target populations, such as workers, with specific exposure risks and for which no robust scientific evidence is currently available. These individuals could definitely benefit from early detection of chronic disease, since the conditions are often worsened by occupational exposure and result in disability and absence from work. Benefits could be further improved by supplementing existing models with information on lifestyle, personal habits and family history (23), social network relationships, income, education and employment history.

FUNDING, CONFLICT OF INTEREST AND AUTHORS' AFFILIATION

The authors confirm that no sources of funding were obtained.
The authors declared no conflict of interest.

DATA AVAILABILITY

No additional data available.

REFERENCES

1 Health statistics and information systems. Disease burden and mortality estimates. CAUSE-SPECIFIC MORTALITY, 2000–2016. Available at: http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html. Last access: 17 January 2019.

2 Mills KT, Bundy JD, Kelly TN et al. Global Disparities of Hypertension Prevalence and Control: A Systematic Analysis of Population-Based Studies From 90 Countries. *Circulation*. 2016 Aug 9;134(6):441-50.doi: 10.1161/CIRCULATIONAHA.115.018912.

3 Global Health Observatory (GHO) data. Raised blood pressure. Situation and trends, Available at: http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence_text/en/. Last access: 18 June 2018.

4 WHO - Health statistics and information systems. DISEASE BURDEN, 2000–2016. Available at: https://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html. Last access: 17 January 2019.

5 Shaw LJ, Goyal A, Mehta C, et al. 10-Year Resource Utilization and Costs for Cardiovascular Care. *J Am Coll Cardiol*. 2018 Mar 13;71(10):1078-1089. doi: 10.1016/j.jacc.2017.12.064.

6 American Diabetes Association. Economic Costs of Diabetes in the U.S. in 2017. *Diabetes Care*. 2018 May;41(5):917-928. doi: 10.2337/dci18-0007. Epub 2018 Mar 22.

7 High Blood Pressure. Blood Pressure: Make Control Your Goal Infographic, Available at: https://www.cdc.gov/bloodpressure/infographic.htm. Last access: 18 June 2018.

8 Mehta S, Ghosh S, Sander S, et al. Differences in All-Cause Health Care Utilization and Costs in a Type 2 Diabetes Mellitus Population with and Without a History of Cardiovascular Disease. *J Manag Care Spec Pharm*. 2018 Mar;24(3):280-290.doi: 10.18553/jmcp.2018.24.3.280.

9 Kung HC, Xu J. Hypertension-related Mortality in the United States, 2000-2013. *NCHS Data Brief*. 2015 Mar;(193):1-8.

10 Gupta R, Gupta VP, Prakash H, et al. 25-Year trends in Hypertension prevalence, awareness, treatment, and control in an Indian urban population: Jaipur Heart Watch. (2018) *Indian Heart Journal*, 70 (6), pp. 802-807.doi: 10.1016/j.ihj.2017.11.011.

11 Bhatnagar P, Wickramasinghe K, Wilkins E, et al. Trends in the epidemiology of cardiovascular disease in the UK. *Heart*. 2016 Dec 15;102(24):1945-1952. doi: 10.1136/heartjnl-2016-309573. Epub 22 August 2016.

12 European Heart Network. European Cardiovascular Disease Statistics 2017, Available at: http://www.ehnheart.org/CvD-statistics/CvD-statistics-2017.html. Last access: 18 June 2018.

13 Petrie JR, Guzik TJ, Touyz RM. Diabetes, Hypertension, and Cardiovascular Disease: Clinical Insights and Vascular Mechanisms. *Can J Cardiol*. 2018 May;34(5):575-584.doi:10.1016/j.cjca.2017.12.005. Epub11 December 2017.

14 Moher D, Liberati A, Tetzlaff J, et al; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009 Jul 21;6(7):e1000097.doi: 10.1371/journal.pmed.1000097. Epub21 July 2009.

15 Sharif MO, Janjua-Sharif FN, Ali H, et al. Systematic reviews explained: AMSTAR-how to tell the good from the bad and the ugly. *Oral Health Dent Manag*. 2013 Mar;12(1):9-16.

16 The Joanna Briggs Institute Reviewers' Manual: 2014 edition / Supplement Methodology for JBI Umbrella Reviews. Available at: http://joannabriggs.org/assets/docs/sumari/ReviewersManual-Methodology-JBI_Umbrella%20Reviews-2014.pdf. Last access: 17 January 2019.

17 Whiting P, Savović J, Higgins JP, et al; ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016 Jan;69:225-34.doi:10.1016/j.jclinepi.2015.06.005. Epub16 June 2015.

18 Abbasi A, Peelen LM, Corpeleijn E, et al. Prediction models for risk of developing type 2 Diabetes: systematic literature search and independent external validation study. *BMJ*. 2012 Sep 18;345:e5900.doi: 10.1136/bmj.e5900.

19 Barber SR, Davies MJ, Khunti K, et al. Risk assessment tools for detecting those with pre-Diabetes: a systematic review. *Diabetes Res Clin Pract*. 2014 Jul;105(1):1-13. doi: 10.1016/j.diabres.2014.03.007. Epub18 March 2014.

20 Collins GS, Mallett S, Omar O, et al. Developing risk prediction models for type 2 Diabetes: a systematic review of methodology and reporting. *BMC Med*. 2011 Sep 8;9:103. doi: 10.1016/j.diabres.2014.03.007.

21 Hu PL, Koh YL, Tan NC. The utility of Diabetes risk score items as predictors of incident type 2 Diabetes in Asian populations: An evidence-based review. *Diabetes Res Clin Pract*. 2016 Dec;122:179-189. doi: 10.1016/j.diabres.2016.10.019. Epub 2 November 2016.

22 Noble D, Mathur R, Dent T, et al. Risk models and scores for type 2 Diabetes: systematic review. *BMJ*. 2011 Nov 28;343:d7163.doi: 10.1136/bmj.d7163.

23 Yoshizawa S, Kodama S, Fujihara K, et al. Utility of nonblood-based risk assessment for predicting type 2 Diabetes mellitus: A meta-analysis. *Prev Med*. 2016 Oct;91:180-187.doi: 10.1016/j.ypmed.2016.07.026. Epub26 July 2016.

24 Cortes-Bergoderi M, Thomas RJ, Albuquerque FN, et al. Validity of cardiovascular risk prediction models in Latin America and among Hispanics in the United States of America: a systematic review. *Rev Panam Salud Publica*. 2012 Aug;32(2):131-9.

25 Tzoulaki I, Liberopoulos G, Ioannidis JP. Assessment of claims of improved prediction beyond the Framingham risk score. *JAMA*. 2009 Dec 2;302(21):2345-52.doi: 10.1001/jama.2009.1757.

26 Fowkes FG, Murray GD, Butcher I, et al. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. *JAMA*. 2008 Jul 9;300(2):197-208.doi: 10.1001/jama.300.2.197.

27 Siontis GC, Tzoulaki I, Siontis KC et al. Comparisons of established risk prediction models for cardiovascular disease: systematic review. *BMJ*. 2012 May 24;344:e3318.doi: 10.1136/bmj.e3318.

28 Damen JA, Hooft L, Schuit E, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ*. 2016 May 16;353:i2416.doi: 10.1136/bmj.i2416.

29 Beswick AD, Brindle P, Fahey T, Ebrahim S. A Systematic Review of Risk Scoring Methods and Clinical Decision Aids Used in the Primary Prevention of Coronary Heart Disease (Supplement) [Internet]. *London: Royal College of General Practitioners (UK); 2008 May. National Institute for Health and Clinical Excellence: Guidance.*

30 Echouffo-Tcheugui JB, Greene SJ, Papadimitriou L, et al. Population risk prediction models for incident heart failure: a systematic review. *Circ Heart Fail*. 2015 May;8(3):438-47.doi:10.1161/CIRCHEARTFAILURE.114.001896. Epub3 March 2015.

31 Damen JA, Pajouheshnia R, Heus P, Moons KGM, Reitsma JB, Scholten RJPM, Hooft L, Debray TPA. Performance of the Framingham risk models and pooled cohort equations for predicting 10-year risk of cardiovascular disease: a systematic review and meta-analysis.*BMC Med.*2019 Jun 13;17(1):109. doi: 10.1186/s12916-019-1340-7.

32 Sun D, Liu J, Xiao L, et al. Recent development of risk-prediction models for incident Hypertension: An updated systematic review. *PLoS One*. 2017 Oct 30;12(10):e0187240. doi: 10.1371/journal.pone.0187240. eCollection 2017.

33 Echouffo-Tcheugui JB, Batty GD, Kivimäki M, et al. Risk models to predict Hypertension: a systematic review. *PLoS One*. 2013 Jul 5;8(7):e67370doi: 10.1371/journal.pone.0067370. Print 2013.

34 Pocock SJ, McCormack V, Gueyffier F, et al. A score for predicting risk of death from cardiovascular disease in adults with raised blood pressure, based on individual patient data from randomised controlled trials. *BMJ*. 2001 Jul 14;323(7304):75-81.

35 Collins DRJ, Tompson AC, Onakpoya IJ, Roberts N, Ward AM, Heneghan CJ. Global cardiovascular risk assessment in the primary prevention of cardiovascular disease in adults: systematic review of systematic reviews. *BMJ Open*. 2017;7:e013650. doi: 10.1136/bmjopen-2016-013650.

36 Studzinski K, Tomasik T, Krzyszton J, Jozwiak J, Windak A. Effect of using cardiovascular risk scoring in routine risk assessment in primary prevention of cardiovascular disease: an overview of systematic reviews. *BMC Cardiovasc Disord*. 2019;19(1):11.

37 Forouzanfar MH, Afshin A, Alexander LT, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the global burden of disease study 2015. The Lancet 2016;388:1659–724. 10.1016/S0140-6736(16)31679-8

38 Kontis V, Mathers CD, Bonita R, et al. Regional contributions of six preventable risk factors to achieving the 25×25 non-communicable disease mortality reduction target: a modelling study. Lancet Glob Health 2015;3:e746–7. 10.1016/S2214-109X(15)00179-5.

39 Loef M, Walach H. The combined effects of healthy lifestyle behaviors on all cause mortality: a systematic review and meta-analysis. Prev Med 2012;55:163–70. 10.1016/j.ypmed.2012.06.017.

40 Holt-Lunstad J, Smith TB, Layton JB, et al. And mortality risk: a meta-analytic review. PLoS Med 2010;7:e1000316.

AUTHORS' CONTRIBUTION

Conceptualization: FL, LP

Methodology: AV, DCM, FL, MM

Project Administration: FL

Visualization: FL, DCM, MM, LP, AA, GA, GB, LM, AV

Writing – Original draft: AA, AV, DCM, FL, MM

Writing – Reviewing and editing: AV, DCM, FL, LP, MM

Validation: FL, LP

Figures Captions

Figure 1 – Study workflow

Figure 2 – Quality assessment scores

Figure 3 – Graphical representation of risk of bias according to the ROBIS tool

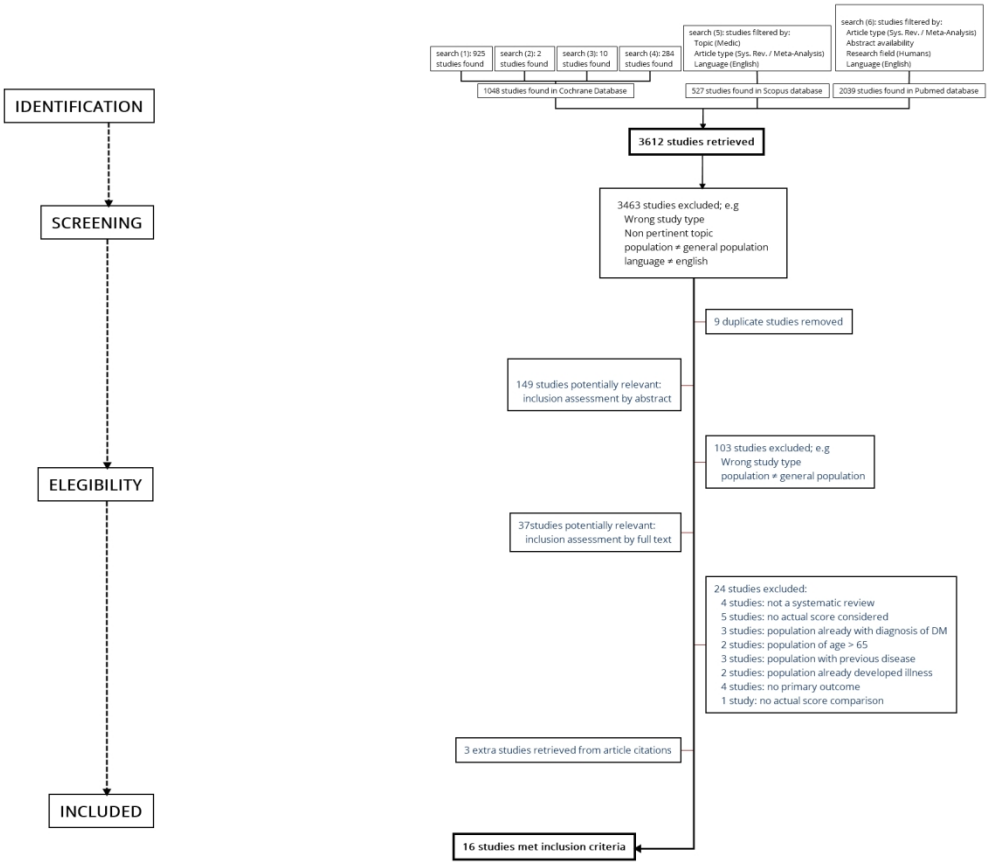

Figure 1 - Study workflow

145x127mm (300 x 300 DPI)

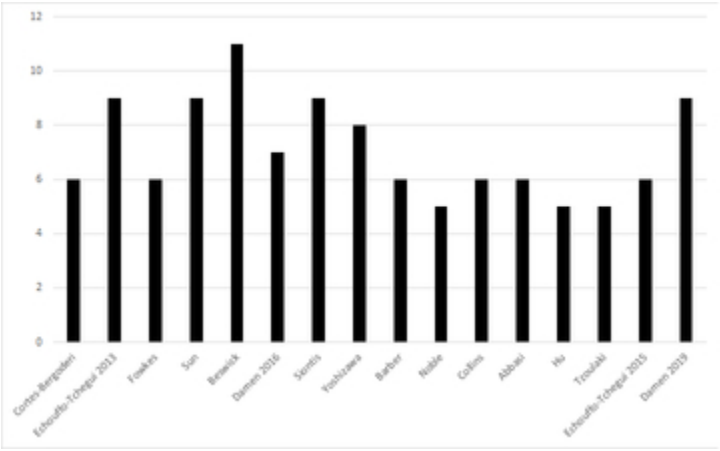

Figure 2 - Quality assessment scores

30x18mm (300 x 300 DPI)

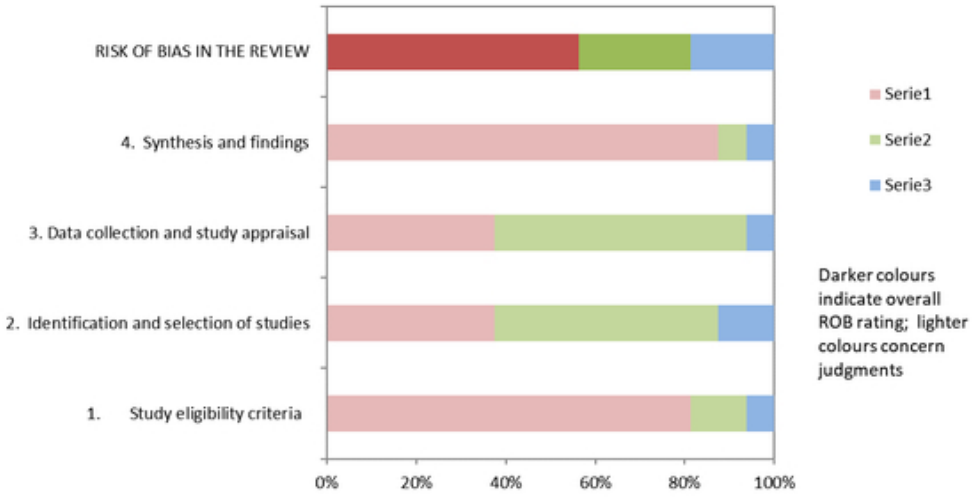

Figure 3 - Graphical representation of risk of bias according to the ROBIS tool

47x25mm (300 x 300 DPI)